(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 416 673 A1**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 90202108.8

(22) Date of filing: 02.08.90

(51) Int. Cl.5: C12N 15/62, C12N 15/48, C12N 15/71, C12P 21/02, G01N 33/68

The microorganism(s) has (have) been deposited with Centraalbureau voor Schimmelcultures, Baarn under number(s) CBS 307.90, CBS 306.90 and CBS 305.90.

(30) Priority: 03.08.89 CU 14989

(43) Date of publication of application: 13.03.91 Bulletin 91/11

(84) Designated Contracting States: AT BE CH DE ES FR GB IT LI NL SE

(71) Applicant: CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA
31 Street, '/156 & 190,
Cubanacan. Playa Havana(CU)

(72) Inventor: Novoa Pérez, Lidia Inés
Calle 184 No.3112, Apto 49, entre 31 y 33
Playa, Ciudad de la Habana(CU)
Inventor: Machado Lahera, Jorge A.
Calle 184 No.3112, Apto 48, entre 31 y 32
Playa, Ciudad de la Habana(CU)
Inventor: Fernández Maso, Julio Raul
Calle 26 No.73, Apto 3, entre 45 y Conill.
Nuevo Vedado, Ciudad de la Habana(CU)
Inventor: Benitez Fuentes, Jes s V.
Calle 184 No.3112, Apto 57, entre 31 y 33
Playa, Ciudad de la Habana(CU)
Inventor: Narciandi Diaz, Ramon Emilio, Ave 391 No.18218
entre 182 y 184, Stgo de las Vegas
Boyeros, Ciudad de la Habana(CU)
Inventor: Rodriquez Reinoso, José Luis
Ave 51 No.1812, entre 180 y 190 Versalles
La Lisa, Ciudad de la Habana(CU)
Inventor: Estrada Garcia, Mario Pablo
Calle 184 No.3112, entre 31 y 33, Apto 50
Playa, Ciudad de la Habana(CU)
Inventor: Garcia Suarez, José
San Rafael No. 1203,entre Mazon y Ronda
Plaza, Ciudad de la Habana(CU)
Inventor: Herrera Martinez, Luis Saturnino
Calle 96 entre 3a y 3a A.
Playa, Ciudad de la Habana(CU)

(74) Representative: Smulders, Theodorus A.H.J., Ir. et al
Vereenigde Octrooibureaux Nieuwe Parklaan 107
NL-2587 BP 's-Gravenhage(NL)

(54) Method for the expression of heterologous proteins produced in fused form in E. coli, use thereof, expression vectors and recombinant strains.

(57) The present invention relates to the field of biotechnology and in particular the use of recombinant DNA technology for the production of heterologous proteins.

The technical object thereof is to develop a highly efficient method for the expression of heterologous genes in fused form in E . coli , which code for proteins which can easily be purified owing to the fact that they are synthesized in insoluble form in the cellular cytoplasm.

To achieve this, an expression vector is used which contains a stabiliser sequence which codes only for the first 58 amino acids belonging to the N-terminal end of the human protein interleukine-2, which is under the tryptophan promoter of the actual E . coli . This vector further contains the gene for resistance to ampicillin as a selection marker and the terminator of transcript ion of bacteriophage T4. In particular the genes which code for the antigenic proteins of the human immunodeficiency virus (HIV 1 and 2) were cloned therein, high levels of expression of said proteins being obtained from transformed strains of the bacteria Escherichia coli .

EP 0 416 673 A1

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 90 20 2108

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
X	EP-A-0 229 998 (HOECHST AG) * Claims 1-3,9 *	1-2	C 12 N 15/62 C 12 N 15/48 C 12 N 15/71 C 12 P 21/02 G 01 N 33/68
X	EP-A-0 227 938 (HOECHST AG) * Claims 1,3,4-7 *	1-2	
X	EP-A-0 227 169 (AKZO N.V.) * Claims 18-20,27-30 *	3,5-9	
X	EP-A-0 345 792 (F. HOFFMANN-LA ROCHE AG) * Claims *	3,5-9	
X	INDIAN JOURNAL OF BIOCHEMISTRY AND BIOPHYSICS, vol. 25, December 1988, pages 504-509; S.G. DEVARE et al.: "Genes of human immunodeficiency virus, type I (HIV-I), their expression in Escherichia coli, and their utility in diagnosis of virus infection" * Whole document *	3,5-9	

TECHNICAL FIELDS SEARCHED (Int. Cl.5)
C 07 K C 12 N C 12 P

The present search report has been drawn up for all claims

Place of search	Date of completion of search	Examiner
The Hague	14 November 90	NAUCHE S.A.

CATEGORY OF CITED DOCUMENTS

X: particularly relevant if taken alone
Y: particularly relevant if combined with another document of the same catagory
A: technological background
O: non-written disclosure
P: intermediate document
T: theory or principle underlying the invention

E: earlier patent document, but published on, or after the filing date
D: document cited in the application
L: document cited for other reasons

&: member of the same patent family, corresponding document

The method which is the subject of the present invention can be employed for the expression at high levels of recombinant heterologous proteins synthesized in fused and insoluble form in E . coli , which can be used in the pharmaceutical industry to obtain vaccine preparations or in the development of diagnostic systems, in the food industry, in agriculture, etc.

METHOD FOR THE EXPRESSION OF HETEROLOGOUS PROTEINS PRODUCED IN FUSED FORM IN E . COLI , USE THEREOF, EXPRESSION VECTORS AND RECOMBINANT STRAINS

The present invention relates to the field of biotechnology and recombinant DNA techniques and in particular to a method for the expression of heterologous proteins synthesized in fused and insoluble form from recombinant E . coli bacteria.

The utility of recombinant DNA technology for producing proteins of interest, of any origin in E . coli , has been extensively demonstrated. For this, a large number of vectors have been developed, although new variants are still necessary owing to the fact that each gene to be cloned and expressed represents an individual case (Denhardt, D.T. and Colasanti, J., Vectors, Butterworths. Stoneham, MA. pp. 179-204, 1987 and Lukacsovich, T. et al., Journal of Biotechnology, 13, 243-250, 1990).

Many eukaryotic polypeptides of clinical or industrial interest, the natural availability of which is scarce, have been obtained by cloning and expression of the genes which code for them in Escherichia coli .

An important problem associated with the production of recombinant proteins in microorganisms is degradation of the product by the host system's own proteases. The stability of the protein can be influenced by different factors such as location of the gene product (Talmadge K. and Gilbert W., Proc. Natl. Acad. Sci. USA 79, 1830-1833, 1982; Moks T. et al., Biochemistry 26, 5239-5244, 1987), selection of the host strain (Buell G. et al., Nucleic Acids Res. 13, 1923-1938, 1985; Bishai W.R. et al., J. Bacteriol. 169, 5140-5151, 1987; Grodberg J. and Dunn, J.J., Bacteriol. 170, 1245-1253, 1988) as well as the conditions of subsequent cultivation and purification (Kitano, K. et al., J. Biotechnol. 5, 77-86, 1987).

Eukaryotic genes cloned in phase with bacterial or synthetic nucleic acid sequences can be expressed as hybrid products in the cellular cytoplasm. Transcription from bacterial promoters as well as translation thereof yields fusion proteins which include bacterial or synthetic polypeptide sequences in addition to the eukaryotic polypeptides (Marston, F.A.O., Biochem. J. 240, 1-12, 1986).

Intracellular synthesis of a fusion protein by expression of a heterologous gene of interest fused to a well-expressed host gene, is a valid means of obtaining high levels of expression of a heterologous protein as well as an increase in stability of the product obtained (Itakura, K. et al., Science, 198, 1056-1063, 1977).

One of the systems used more for this purpose has been to obtain proteins fused to the beta-galactosidase from E . coli (Itakura, K. et al., Science, 198, 1056-1063, 1977). However, the main disadvantage of this system is the large size of this protein, on account of which the desired peptide represents only a small portion of the total hybrid protein (Flores, N. et al., Appl. Microbiol. Biotechnol. 25, 267-271, 1986; Goeddel, D.V. et al., PNAS USA, 76: 106-110).

German patent no. 35 41 856 A1 (Hoechst AG) reports on the possibility of using a stabiliser peptide consisting of at least the first 95 amino acids of the N-terminal end of the human protein interleukine-2 to obtain fusion proteins in insoluble form synthesized in E . coli , with a view to expressing eukaryotic peptides such as proinsulin and hirudine, without reference to the levels of expression reached with this system. In this patent are also included in the genetic construction particular sequences for cleavage of the end product with a view to separating the protein of interest from the stabiliser peptide.

The production of viral proteins by genetic engineering is of great interest for the development of methods of diagnosis and vaccine preparations, above all because of the purity of the resulting products as well as the elimination of manipulation of the active pathogenic agent. In the field of diagnosis, these products are of great importance in early detection of antibodies to these organisms, high specificity and sensitivity in said systems being achieved.

In particular, in the case of human retroviruses, it is necessary to develop highly sensitive systems for the detection of antibodies on the basis of very pure antigens, avoiding any loss of specificity which would invalidate the use of them. These organisms cause various immunological changes, depending on the particular subgroup to which the viral agent belongs, and also due to its trophism for T-lymphocyte cells, being able to cause abnormal proliferation or impaired functionality of said cells (leukaemia) or a depletion of the cell population (immunosuppression) (Wong-Staal, F. and Gallo, R.C., Nature, 317, 395-403, 1985).

It is therefore necessary to count on efficient systems of expression of the main proteins with antigenic activity belonging to the viruses which cause these diseases, with a view to using them in rapid and precise diagnostic systems, which will make it possible to carry out large-scale epidemiological studies for the detection of antibodies to these viruses during processing of blood samples in banks and thus to prevent the disease from being transmitted by this pathway.

The genes which code for the main proteins with antigenic activity of human immunodeficiency viruses (HIV) have been cloned and expressed in E . coli , both directly and fused to other genes belonging to said host.

Among the proteins expressed in their natural form are peptide 121 of AIDS, which is obtained in insoluble form with levels of expression varying between 5 and 10% of total proteins (Chang, T.W. et al., Biotechnology 3, 905-909, 1985) and protein gag 24 of the same virus which is obtained in soluble form at levels not calculated (Dowbenko, D.J. et al., Proc. Natl. Acad. Sci. USA, 82, 7748-7752, 1985).

In Spanish patent no. 2 000 859 (Syntex) is described a method for the expression of fusion proteins using a vector which contains a DNA gene of the protein TrpLE of E . coli , in which is specifically inserted a DNA sequence of the AIDS virus. In this case, the carboxy-terminal LE region is substituted by a heterologous polypeptide, as a result of which a self-aggregating fusion protein is obtained, purification thereof being simplified in this way. Moreover the vector used contains binding means for three reading frames which facilitates isolation of the protein of interest. In this patent is described the construction of a clone of high expression which produces more than 5% of the total cell protein.

The present invention relates to a method for the expression of heterologous proteins produced in insoluble form in E . coli and in particular fusion proteins which contain a fragment or the whole of a viral protein such as the case of antigenic proteins belonging to human immunodeficiency virus (HIV 1 and 2). For this, there was used a vector which contains a stabiliser sequence wich codes for approximately the first 58 amino acids of the N-terminal end of human interleukine-2 (IL-2), which guarantees high levels of expression of the heterologous genes cloned. This vector further consists of the tryptophan promoter of E . coli (ptrp), the gene for resistance to ampicillin as a selection marker, the terminator of transcription of bacteriophage T4 and restriction sites Xba I, Xho I and Bam HI for coupling of the genes which it is desired to express. The present invention therefore also relates to the expression vectors used for cloning and expression of the different antigenic proteins of HIV 1 and 2 in E . coli as well as the recombinant strains obtained, which express levels of said heterologous proteins varying between 20 and 25% of the total proteins produced by them.

In particular, the proteins expressed were the one belonging to the nucleus (gag24) and a fragment of the coat protein (gp41) of virus HIV 1 and a fragment of the transmembraneous protein gp36 of HIV 2. The strains used as hosts for cloning of the genes which code for these proteins were E . coli K-12 HB-101, W-3110 and C-600 respectively.

An innovating feature of the present invention is the use of a stabiliser sequence, which consists of a fragment of the N-terminal end of the gene of human interleukine-2 protein which codes only for the first 58 amino acids of said protein, which is used for the expression of heterologous proteins and in particular the main proteins with antigenic activity of the HIV virus.

The fusion proteins expressed by means of the method described are synthesized in insoluble form, which simplifies the final purification process and makes it more efficient, on account of which proteins which display antigenic activity are obtained, which are used in diagnostic methods for the detection of antibodies to them without the need for cleavage of the stabiliser fragment used in the fusion, the present invention also relating to use of the fusion protein obtained.

EXAMPLES

Example 1

For the expression of different heterologous proteins in E . coli , there was constructed the expression vector pFP-15, in which was inserted the sequence which codes for a stabiliser peptide, consisting of the first 58 amino acids belonging to the N-terminal fragment of the protein of human origin, interleukine-2 (IL-2) . Said sequence is cloned under the control of the tryptophan promoter of E . coli , said vector further comprising the terminator of bacteriophage T4 as a signal of termination of transcription and the gene for ampicillin resistance as a selection marker.

The plasmid vector pFP-15 was constructed by ligation of a synthetic oligonucleotide of 190 bases and its complementary one, which contains the sequence which codes for the first 58 amino acids of the N-terminal end of IL-2 (Fig. 1), the stabiliser sequence, and the vector pTPV-1 (Fig. 2) which carries the tryptophan promoter of Escherichia coli and the terminator of bacteriophage T4. The layout of said construction is shown in Fig. 2.

Coupling of the DNA segment which codes for the above stabiliser peptide was verified by DNA sequence analysis according to the description in the literature (Sanger, F. et al., PNAS, USA, 74, 5463-5467, 1977) using an oligonucleotide (Fig. 3) which hybridises with the ptrp promoter and sequence in the direction of the stabiliser (5'-3'). Thus it was possible to check that in all cases the appropriate reading

frame was maintained.

Example 2

For cloning and expression of the nuclear protein of virus HIV-1 (gag24), the following oligonucleotides were designed:

5′ CAT CTA GAC ATG CAA ATG TTA AAA GAA 3′

3′ GT TTA GGT CGA TTG ACT ATC CTA GGC 5′

These oligonucleotides correspond to the 5′ and 3′ ends respectively of the gene which codes for protein gag24 (Alizon, M. et al., Nature 312, 757-760, 1984). With these oligonucleotides and with the genome of HIV-1 isolated, amplification was carried out by the technique of the polymerase chain reaction (PCR) (Randall, K. et al., Science, USA, 239, 487-491, 1988) of the gene which codes for a fragment of the gag24 gene. This fragment was cut at sites Xba I and Bam HI, which were contained in the oligonucleotides used in the PCR, and was ligated to the expression vector pFP-15, Xba I-Bam HI being digested, the amplified gene being thus ligated to the segment which codes for the stabiliser peptide, under the tryptophan promoter. The recombinant plasmid obtained, called VIHCA (Fig. 4), was transformed in cells of E . coli strain K-12 HB-101. The transformed colonies were selected for ampicillin resistance in dishes of Luria broth medium (Miller, J.H., Cold Spring Harbor Lab., 1972) supplemented with the antibiotic at 50 ug/ml final concentration, and the recombinants were identified by the technique of hybridisation, using as a radioactive probe (labeled with ^{32}P) the actual amplified fragment used for cloning. An immunoidentification test was carried out on the positive ones in autoradiography, with serum of infected patients and ^{125}I labeled protein A, expression of the protein gag24 being identified by the positivity of these clones in the immunological technique. On these individuals was carried out the Western blot technique (Burnette, W.N., Anal. Bioch., 112, 195-203, 1981), a band of approximately 28,000 daltons being obtained, which corresponds to the length of the stabiliser peptide (approximately 58 amino acids) plus the fragment of the cloned protein gag24 (approximately 180 amino acids).

Example 3

For cloning and expression of the transmembraneous protein gp41, first of all the synthesis of an oligonucleotide of 269 bases and its complementary one (Fig. 5) was carried out, corresponding to a fragment of said protein belonging to the coat of the virus (Han, B.H. et al., Nature 312, 166-169, 1984). This oligonucleotide was digested with Bam HI and ligated to the vector pPF-15 previously cut with Xba I, treated with S1 nuclease and finally digested by Bam HI, the desired gene remaining fused to the segment which codes for the stabiliser peptide, under the tryptophan promoter of said vector. The product of this ligation is the vector VIHTA-1 (Fig. 6) with which E . coli strain K-12 W-3110 was transformed. The transformed colonies were selected for ampicillin resistance, and the recombinants were identified by the technique of hybridisation, an immunoidentification test being performed on the positive ones as in the preceding example. Western blot was carried out on one of the individuals which showed expression of the fused protein, a band of approximately 15,000 daltons being obtained, which corresponds to the expected size of the fusion protein, which includes the 58 amino acids of the stabiliser peptide plus 83 amino acids corresponding to the fragment of protein gp41 of HIV-1.

Example 4

Cloning was carried out of the region representing the gene which codes for expression of the transmembraneous protein of HIV-2, gp36, by the synthesis of an oligonucleotide of 318 bp and its complementary one (Fig. 7) corresponding to a fragment of the protein gp36 of the coat of HIV-2 (Clavel, F. et al., Science 233, 343-346). This DNA segment was ligated to the vector pPF-15 previously cut by Xba I/ Bam HI, the desired fragment remaining fused to the gene which codes for the stabiliser peptide, under the tryptophan promoter of said vector. The product of this ligation is the vector VIHTA-2 (Fig. 8), which was inserted in E . coli strain K-12 C-600. The transformed colonies were selected for ampicillin resistance, and the recombinants were identified by the technique of hybridisation, an immunoidentification test being performed on the positive ones as in examples 2 and 3.

In all cases, coupling of the DNA segments to the stabiliser was verified by DNA sequence analysis, as

reported in the literature (Sanger, F. et al., PNAS, USA, 74, 5463-5467, 1977).

Example 5

In the case of fusion proteins gag24-stabiliser peptide and gp41-stabiliser peptide, the respective recombinant strains W41 were grown in super broth medium (32 g tryptone and 20 g yeast extract per litre of distilled water) supplemented with $FeCl_3$ (0.001 mM), $MgSO_4$ (0.1 mM), M9 salts (N at 6%, KH_2PO_4 at 3%, NaCl at 0.5% and NH_4Cl at 1%) and ampicillin 50 ug/ml.

For protein gp36-stabiliser peptide, the transformed strain C36 was grown on minimal medium (M al., Cold Spring Harbor Lab., USA, 1982) supplemented with casein hydrolysate at 2%, glucose at 2 $MgSO_4$, 0.1 mM $CaCl_2$ and ampicillin at the same concentration as in the previous case.

of the cultures was carried out at an optical density of 0.05, maintaining them at 37° C for 12 hours, with agitation at 260 rpm, aeration at 1 vvm, finally reaching an optical density of 10 read at 600 nm, induced by depletion of the tryptophan by the addition of indoleacrylic acid (Squires, C.L. et al., Jour. of Mol. Biol., USA, 92, 93-111, 1975) two hours after the start of fermentation. The cells obtained are collected by centrifuging and stored at -20° C to be used subsequently in recovery of the desired product. After ultrasonic rupture of the biomass, levels of expression of 20-25% of the total protein are ascertained by SDS-PAGE electrophoresis of proteins (Laemmli, Nature, UK, 227, 680-685, 1970) and analysis of the series on a SCANNER 65 300, USA.

STRAIN DEPOSITS

The E . coli HB24 [pVIHCA] strain, based on the E . coli strain K-12 HB-101 and containing the plasmid pVIHCA, was deposited on July 11, 1990, with the Centraalbureau voor Schimmelcultures (CBS), Baarn, The Netherlands, and obtained deposit number CBS....90.

The E . coli W41 [pVIHTA-1] strain, based on the E . coli strain K-12 W-3110 and containing the plasmid pVIHTA-1, was deposited on July 11, 1990, with the Centraalbureau voor Schimmelcultures (CBS), Baarn, The Netherlands, and obtained deposit number CBS....90.

The E . coli C36 [pVIHTA-2] strain, based on the E . coli strain K-12 C-600 and containing the plasmid pVIHTA-2, was deposited on July 11, 1990, with the Centraalbureau voor Schimmelcultures (CBS), Baarn, The Netherlands, and obtained deposit number CBS....90.

Sequence Listing

SEQ ID NO:1

SEQUENCE TYPE: Nucleotide with corresponding protein
SEQUENCE LENGTH: 194 base pairs
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE ORGANISM: Human interleukin 2
IMMEDIATE EXPERIMENTAL SOURCE: Nucleotide synthesis
FEATURES: from 8 to 181 bp mature peptide
PROPERTIES: Coding gene for stabilizer peptide

```
CGATTCC ATG GCG CCT ACT TCA AGT TCT ACA AAG AAA ACA   40
        Met Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr
                            5                  10

CAG CTA CAA CTG GAG CAT TTA CTG CTG GAT TTA CAG ATG   79
Gln Leu GLn Leu Glu His Leu Leu Leu Asp Leu Gln Met
            15                  20

ATT TTG AAT GGA ATT AAT AAT TAC AAG AAT CCC AAA CTC  118
Ile Leu Asn Gly Ile Asn Asn Tyr Lys Asn Pro Lys Leu
 25                  30                  35

ACC AGG ATG CTC ACA TTT AAG TTT TAC ATG CCC AAG AAG  157
Thr Arg Met Leu Thr Phe Lys Phe Tyr Met Pro Lys Lys
        40                  45                  50

GCC ACA GAA CTG AAA CAT CTC CAG TGTCTAGAGC TAG       194
Ala Thr Glu Leu Lys His Leu Gln
                50
```

SEQ ID NO:2

SEQUENCE TYPE: Nucleotide
SEQUENCE LENGTH: 17 base pairs
MOLECULE TYPE: DNA

TCGAACTAGT TAACTAG 17

SEQ ID NO:3

SEQUENCE TYPE: Nucleotide
SEQUENCE LENGTH: 27 base pairs
MOLECULE TYPE: DNA

CATCTAGACA TGCAAATGTT AAAAGAA 27

SEQ ID NO:4

SEQUENCE TYPE: Nucleotide
SEQUENCE LENGTH: 26 base pairs
MOLECULE TYPE: DNA

CGGATCCTAT CAGTTAGCTG GATTTG 26

SEQ ID NO:5

SEQUENCE TYPE: Nucleotide
SEQUENCE LENGTH: 268 base pairs
MOLECULE TYPE: DNA

GGGGAAGCTC AACAACACTT GTTGCAATTG ACTGTTTGGG GTATCAAGCA 50
ATTGCAAGCT AGAATCTTGG CTGTTGAAAG ATACTTGAAG GACCAACAAT 100
TGTTGGGTAT CTGGGGTTGT TCTGGTAAGT TGATCTGTAC TACTGCTGTT 150
CAATGGAACG CTTCTTGGTC TAACAAGTCT TTGGAACAAA TCTGGAACAA 200
CATGACTTGG ATGGAATGGG ACAGAGAAAT CAACAACTAC ACTTCTTTGT 250
AATAGGGATC CGTCGACC 268

SEQ ID NO:6

SEQUENCE TYPE: Nucleotide
SEQUENCE LENGTH: 321 base pairs
MOLECULE TYPE: DNA

```
CTAGAAGTTC AGCAACAACA ACAGTTATTG GACGTAGTTA AGAGACAACA     50
GGAACTATTG AGACTAACCG TTTGGGGAAC CAAGAACTTA CAGGCAAGAG    100
TAACTGCTAT CGAGAAATAT CTACAAGACC AGGCTCGTCT AAATTCATGG    150
GGATGTGCAT TCCGTCAGGT ATGTCACACT ACCGTACCAT GGGTTAATGA    200
TTCTTTAGCT CCAGACTGGG ATAATATGAC CTGGCAGGAG TGGGAAAAGC    250
AAGTACGTTA CTTAGAGGCT AACATTTCAA AAAGTTTGGA GCAGGCACAG    300
ATCCAGGGTA CTAATAGCTA G                                   321
```

SEQ ID NO:7

SEQUENCE TYPE: N-terminal fragment of human interleukin 2
SEQUENCE LENGTH: 58 amino acids
MOLECULE TYPE: Peptide
ORIGINAL SOURCE ORGANISM: Human interleukin 2
IMMEDIATE EXPERIMENTAL SOURCE: Nucleotide synthesis
FEATURES: from 1 to 58 amino acid mature peptide
PROPERTIES: Stabilizer peptide

```
Met Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr Gln
                  5                  10
Leu GLn Leu Glu His Leu Leu Leu Asp Leu Gln Met
         15                  20
Ile Leu Asn Gly Ile Asn Asn Tyr Lys Asn Pro Lys
 25                  30                  35
Leu Thr Arg Met Leu Thr Phe Lys Phe Tyr Met Pro
             40                  45
Lys Lys Ala Thr Glu Leu Lys His Leu Gln
     50                  55
```

Claims

1. A method for the expression of heterologous proteins produced in fused form in E. in which a stabiliser sequence is used for expression of the heterologous proteins, consisting of an N-terminal fragment of human interleukine-2, characterised in that said sequence codes for not more than the first 58 amino acids of this protein, to which is fused the sequence of the heterologous protein to be expressed.

2. A method according to claim 1, characterised in that the amino acid sequence of the stabiliser peptide corresponds to:

```
          10         20         30         40         50
MAPTSSSTKK TQLQLEHLLL DLQMILNGIN NYKNPKLTRM LTFKFYMPKK ATELKHLQ
```

3. A method according to claim 1, characterised in that the heterologous proteins which are expressed correspond to the nuclear protein (gag24) and the transmembraneous protein (gp41) belonging to human immunodeficiency virus HIV-1 and the transmembraneous protein gp36 belonging to human immunodeficiency virus HIV-2.

4. Expression vector pFP-15, characterised in that it contains the stabiliser sequence which codes for tee first 58 amino acids of human interleukine-2 under the tryptophan promoter of E . coli with the signal for termination of bacteriophage T4 and the gene for ampicillin resistance, and contains the restriction sites Xba I, Bam HI and Xho I for fusion of the heterologous protein which is to be expressed.

5. Vectors VIHCA, VIHTA-1 and VIHTA-2 derived from pFP-15, characterised in that they contain gene sequences coding for protein gag24, for a fragment of protein gp4l (both of HIV-1) and for a fragment of protein gp36 of HIV-2 respectively, which are coupled to the stabiliser sequence of vector pFP-15 using the restriction sites present therein.

6. Recombinant strains HB24, W41 and C36, characterised in that they are obtained as a result of transformation of E . coli strains K-12 HB-101, W-3110 and C-600 with the vectors VIHCA, VIHTA-1 and VIHTA-2 respectively and that they express high levels of the antigenic HIV proteins in insoluble form.

7. Fusion proteins obtained according to the preceding claims, characterised in that they are composed of a peptide which includes the first 58 amino acids belonging to the N-terminal end of human interleukine-2, which is fused to a heterologous protein.

8. Fusion proteins according to claim 7, characterised in that said heterologous protein corresponds to the protein gag24, protein gp41 (both of HIV-1) or protein gp36 of HIV-2.

9. Use of the fusion proteins obtained according to the preceding claims, characterised in that they can be used in diagnostic methods for the detection of human or animal antibodies.

```
1
CGATTCCATG GCGCCTACTT CAAGTTCTAC AAAGAAAACA CAGCTACAAC TGGAGCATTT
61
ACTGCTGGAT TTACAGATGA TTTTGAATGG AATTAATAAT TACAAGAATC CCAAACTCAC
121
CAGGATGCTC ACATTTAAGT TTTACATGCC CAAGAAGGCC ACAGAACTGA AACATCTCCA
181
GTGTCTAGAG/ctag
```

-Extension CG at 5' end sticky Cla I

-Extension CTAG at 5' end of the complementary strand sticky BamH I

-ATG initiation of transcription

FIG.1

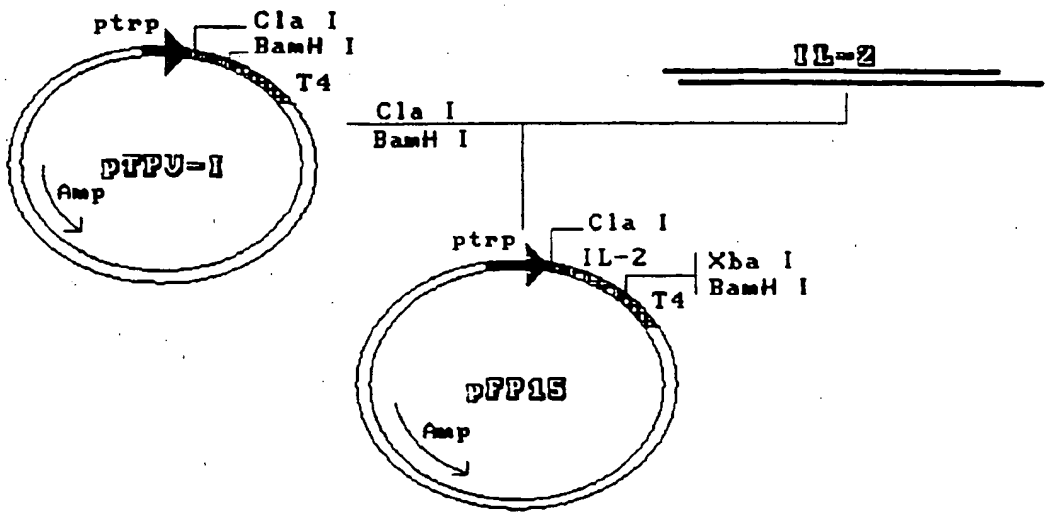
ptrp
Cla I
BamH I
T4
pTPV-I
Amp
IL-2
Cla I
BamH I
ptrp
Cla I
IL-2
Xba I
BamH I
T4
pFP15
Amp

FIG.2

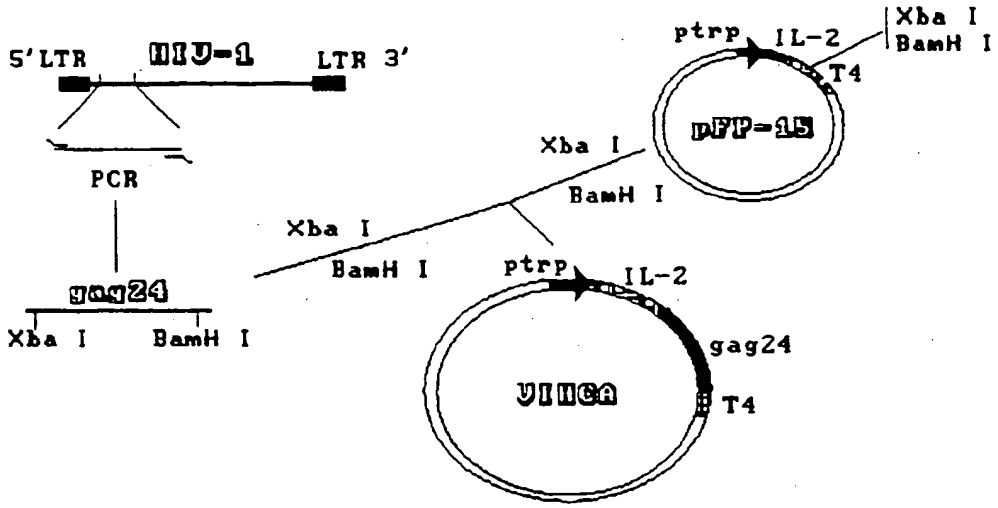
5'LTR
HIV-1
LTR 3'
PCR
gag24
Xba I
BamH I
ptrp
IL-2
T4
Xba I
BamH I
pFP-15
Xba I
BamH I
ptrp
IL-2
gag24
T4
VIHCA

FIG.4

5' TCGAACTAGTTAACTAG 3'

FIG.3

```
1
GGGGAAGCTC AACAACACTT GTTGCAATTG ACTGTTTGGG GTATCAAGCA ATTGCAAGCT
61
AGAATCTTGG CTGTTGAAAG ATACTTGAAG GACCAACAAT TGTTGGGTAT CTGGGGTTGT
121
TCTGGTAAGT TGATCTGTAC TACTGCTGTT CAATGGAACG CTTCTTGGTC TAACAAGTCT
181
TTGGAACAAA TCTGGAACAA CATGACTTGG ATGGAATGGG ACAGAGAAAT CAACAACTAC
241
ACTTCTTTGT AATAGGGATC CGTCGACC
                  BamH I
```

FIG.5

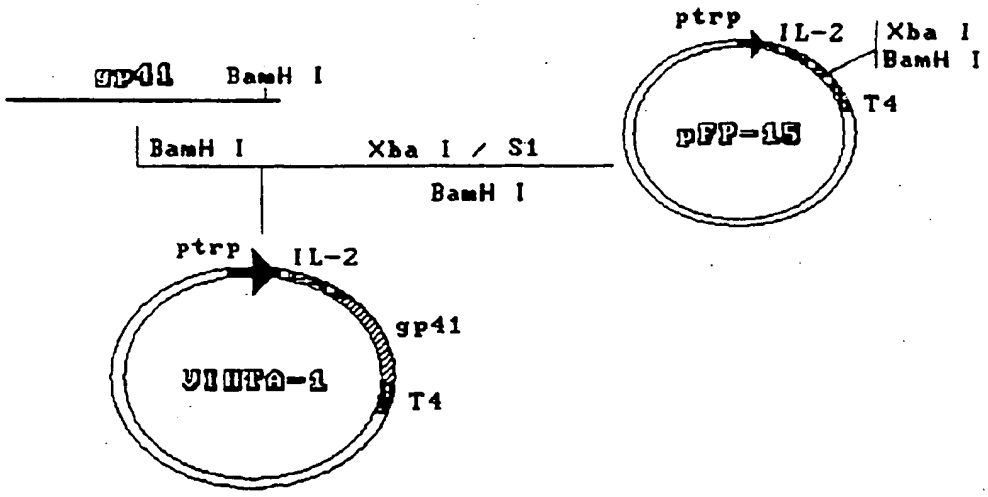
gp41
BamH I
BamH I
Xba I / S1
BamH I
ptrp
IL-2
gp41
T4
VIHTA-1
ptrp
IL-2
Xba I
BamH I
T4
pFP-15

FIG.6

```
1
CTAGAAGTTC AGCAACAACA ACAGTTATTG GACGTAGTTA AGAGACAACA GGAACTATTG
61
AGACTAACCG TTTGGGGAAC CAAGAACTTA CAGGCAAGAG TAACTGCTAT CGAGAAATAT
121
CTACAAGACC AGGCTCGTCT AAATTCATGG GGATGTGCAT TCCGTCAGGT ATGTCACACT
181
ACCGTACCAT GGGTTAATGA TTCTTTAGCT CCAGACTGGG ATAATATGAC CTGGCAGGAG
241
TGGGAAAAGC AAGTACGTTA CTTAGAGGCT AACATTTCAA AAAGTTTGGA GCAGGCACAG
301
ATCCAGGGTA CTAATAG/ctag
```

-Extension CTAG at 5' end sticky Xba I

-Extension CTAG at 5' end of the complementary strand sticky BamH I

FIG.7

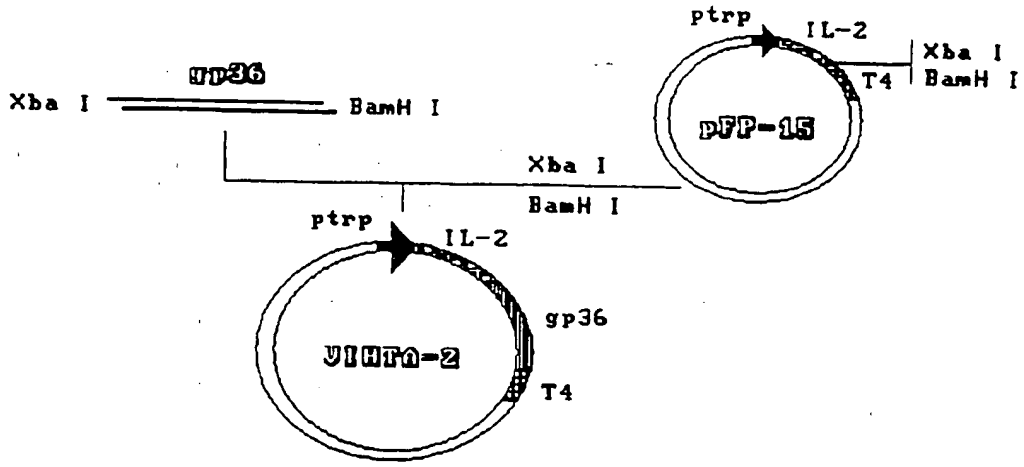
ptrp
IL-2
Xba I
BamH I
T4
pFP-15
gp36
Xba I
BamH I
Xba I
BamH I
ptrp
IL-2
gp36
pIHTa-2
T4

FIG.8